Using Learning Analytics to Detect Off-Task Reading Behaviors in Class

Akçapnar, Gökhan; Hasnine, Mohammad Nehal; Majumdar, Rwitajit; Flanagan, Brendan; Ogata, Hiroaki

Akçapnar, Gökhan ...[et al]. Using Learning Analytics to Detect Off-Task Reading Behaviors in Class. Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge (LAK'19) 2019: 471-476

Using Learning Analytics to Detect Off-Task Reading Behaviors in Class

Gökhan Akçapınar[1,2], Mohammad Nehal Hasnine[1], Rwitajit Majumdar[1], Brendan Flanagan[1], Hiroaki Ogata[1]
Kyoto University[1], Hacettepe University[2]
akcapinar.gokhan.2m@kyoto-u.ac.jp

ABSTRACT: In this paper, we aimed at detecting off-task behaviors of the students by analyzing logs from a digital textbook reader. We analyzed 47 students' reading logs from a 60-minutes long in-class reading activity. During the preprocess, we extracted each student's reading patterns as a single vector. Then we used cluster analysis to find the most common reading patterns. Our results indicated that there are two major reading patterns in data. The first pattern is, the students who are following the instructor from the beginning until the end of the lecture. The second pattern is, students who are following the instructor's pattern until the first 17th minute but not during the rest of the lecture. Based on these patterns we labeled first group as *on-task* students while the other group as *off-task* students. We also investigated academic performance of students in these two groups. Obtained results can be used to design data-driven support for in-class teaching. Instructors can plan interventions when off-task behaviors occur while the lecture is in progress.

Keywords: learning analytics, educational data mining, in-class decision making, off-task behavior, reading pattern analysis, clustering

1 INTRODUCTION

Off-task behaviors can be defined as any actions that a student exhibit in the learning environment that are not according to the tasks given by the lecturer (McElroy-Yeider & Courtney, 2016). Off-task behavior is a common problem that intelligent tutoring systems and traditional classrooms often face (Hughes, 2010). According to Hofer (2007), there are two types of off-task behaviors in traditional classrooms, that are: active and passive. Active off-task behaviors include physical activities that students exhibit in a learning environment which often considered to be distributing to their surroundings and consequently effects teaching process negatively (e.g. disturbing other students, making noise, etc.). On the other hand, passive off-task behavior means that students are cognitively disengaged from ongoing learning activities (e.g. daydreaming, texting to other students etc.). Passive off-task behaviors may be harder to notice since students are not disturbing their surroundings (McElroy-Yeider & Courtney, 2016).

With regard to online learning environments, abovementioned problems remain when technology is used to support in-class learning. In addition, devices like computers, mobiles, tablets etc. can be a reason of distraction because students may play games, use other applications, and browse internet (Hughes, 2010).

Both active and passive off-task behaviors require teachers' attention that can lead to frustration for teachers and limit the learning scopes within a classroom (Hofer, 2007). Engaging with off-task behaviors has also been shown to be associated with poor learning (Baker, Corbett, Koedinger, & Wagner, 2004; Cocea, Hershkovitz, & Baker, 2009). Therefore, both traditional classrooms and online learning environments should consider reducing off-task behaviors while promoting on-task behaviors.

Previous researches have focused on developing detectors for off-task behaviors for intelligent tutoring systems (Cetintas, Si, Xin, & Hord, 2010; Walonoski & Heffernan, 2006). However, without using biological sensors such as eye trackers or EEG headsets, detecting off-task behaviors in traditional classrooms is challenging task (Baker, 2007). In this paper, we aimed at detecting passive off-task behaviors in classroom setting by using students' reading logs that were collected from a digital textbook reader.

2 METHOD

2.1 Data

As the data source, we used reading logs collected from a 60-minutes long in-class activity. In the class, there were 47 students. Both students and instructor used the digital textbook reader (BookRoll) during the lecture. BookRoll is a system that allows to view digital materials use for delivering lecture (Ogata et al., 2018). It is an online environment that allows teachers to upload contents as pdf file. Students can browse anytime and anywhere from web browser in their personal devices (computer or smartphone).

In the BookRoll system, there are features like bookmark, markers, memo function etc. that students can use for learning. In the collection of data for this study, students used their mobile devices or laptops to access the BookRoll system. Reading logs collected automatically by the learning analytics system developed by Flanagan and Ogata (2017). After 60 minutes learning session, 4430 rows of click-stream were recorded in database that are related to students' interaction with the system. At the end of the lecture, students took part in the quiz session related to content.

2.2 Preprocess

The collected click-stream data contained the following fields: *userid* (anonymized student userid), *contentsid* (the id of the e-book that is being read), *operationname* (the action that was done, e.g. open, close, next, previous, jump, add marker, add bookmark, etc.), *pageno* (the current page where the action was performed), marker (the reason for the marker added to a page, e.g. important, difficult), *memo_length* (the length of the memo that was written on the page), *devicecode* (type of device used to view BookRoll, e.g. mobile, pc), and *eventtime* (the timestamp of when the event occurred). For the analysis, we used *eventtime* and *pageno* columns. We grouped the data into 1-minute intervals, and extracted the pages for each student for all time intervals. If student does not have log for the specific time interval, we assumed that students are in the same page where s/he was in the last time.

2.3 Data Analysis

For the data analysis, first we visualized all students' reading patterns. Later, we calculated relative reading patterns of all students. To do this we took instructor's reading pattern as a baseline since expected reading behavior of students is to follow the instructor during the lecture. Finally, to find off-task reading behaviors we used cluster analysis. Since we do not have prior knowledge about the number of clusters in the data, we conducted GAP statistics (Hastie, Tibshirani, & Walther, 2001) to find the optimal number of clusters. Students took part in the open-book quiz during the last 15 minutes of the lecture, therefore reading patterns of the students during this time is varying. We eliminated quiz part and limited our cluster analysis with the first 45 minute of the course.

3 RESULTS

Visualization of reading patterns of all students can be seen in Fig.1. In the Fig.1, X-axis shows the time, Y-axis shows the page of the books. Intersection of the Time and Page shows the current page of the student in a specific time. Each line shows reading patterns of the different students. Expected reading pattern is, to increase the number of page as the time progresses. As observed from the Fig.1, while most of the students are following the this expected pattern, there are some students who has different reading patterns.

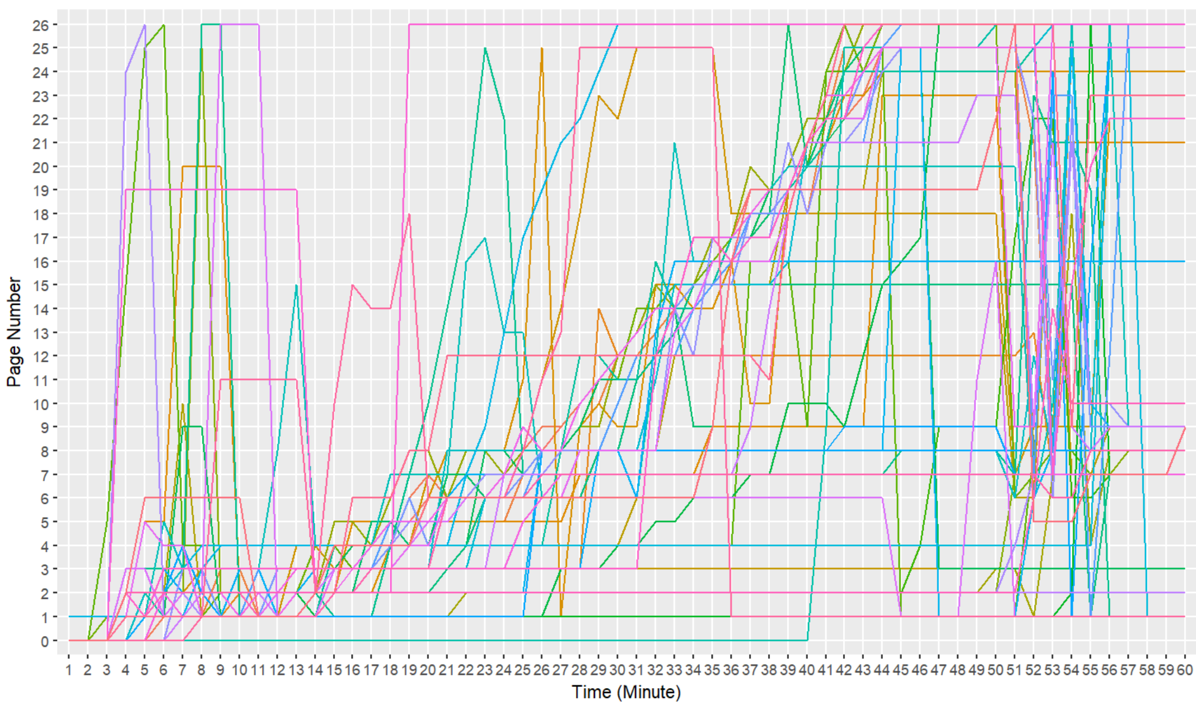

Figure 1: Students' reading patterns across the lecture

To standardize students reading patterns, we calculated relative reading patterns. For instance, if a student is in page 4 while the instructor is in page 6, that student's relative distance will be -2. If a student is in page 8 that student's relative distance will be +2. And if the student is in page 4 (same page as instructor) it will be 0. Results of this calculation is shown in Fig.2. Here again X-axis shows the time of the lecture, while Y-axis shows the students relative distance from the page where instructor is currently in. After calculating students' relative distances from the instructor's pattern, we conducted cluster analysis to find the common reading patterns.

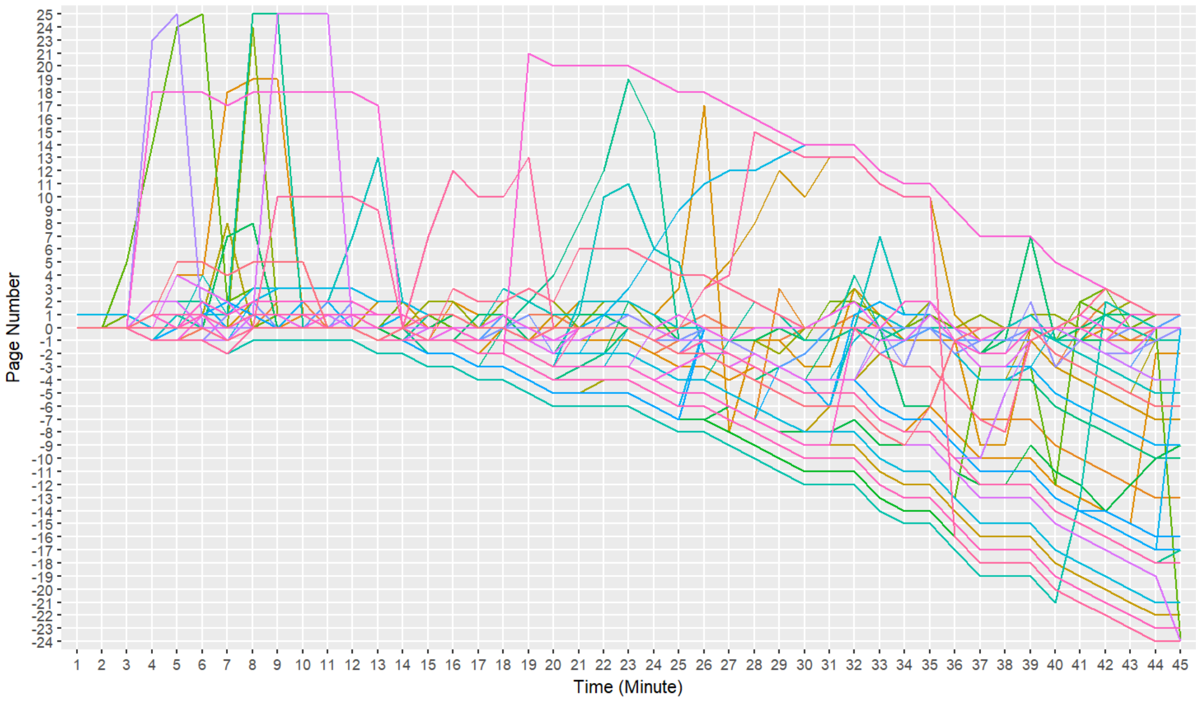

Figure 2: Relative reading patterns of students

3.1 Cluster Analysis Results

Results of the GAP statistics can be seen in Fig.3 (left). According to the results, optimal number of cluster was found as 2. Fig.3 (right) shows the visualization of these two clusters.

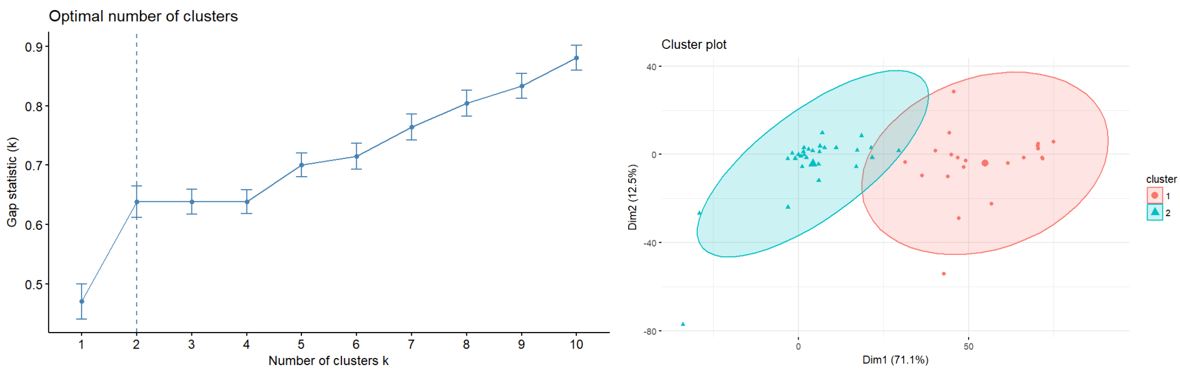

Figure 3: GAP Statistics (left) – Cluster Centers (right)

To see the common reading patterns of the students in these two clusters, we visualized cluster means as well. Results can be seen in Fig.4. From Fig.4, we found two different patterns. Based on these patterns, students in Cluster 2 labelled as On-Task students since they are following the instructor until at the end of the lecture. On the other hand, students in Cluster 1 labelled as Off-Task students since after 17th minute of the lecture those students could not follow the instructor. In addition, distance between Cluster 2 and Cluster 1 increased towards the end of the course.

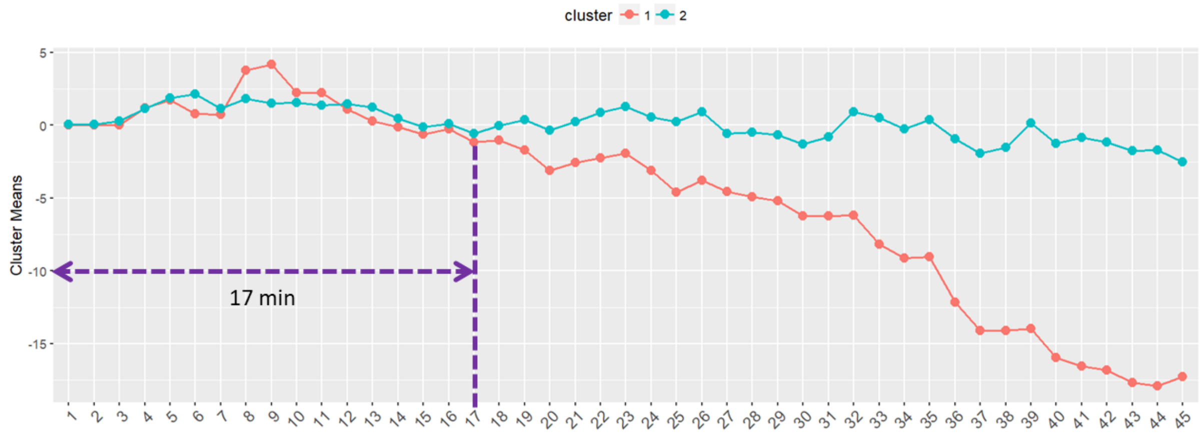

Figure 4: Clustered reading patterns

3.2 Student Academic Performance

As mentioned before students took the quiz in last 15 minutes of lecture. Even it was an open-book quiz, we compared the quiz performance of the students in two different clusters. Since the data were not normally distributed, we used Wilcoxon Signed-ranks test to compare two groups. We compared students' quiz scores and the time they spend on quiz. Results is shown in Table 1. In terms of scores, a Wilcoxon Signed-ranks test indicated no significant difference between Cluster 1 (Mdn = 10) and Cluster 2 (Mdn = 10), W = 296, p = .58. The time spent on quiz was also not significantly differed among Cluster 1 (Mdn = 146) and Cluster 2 (Mdn = 120), W = 320, p = .31.

Table 1: Descriptive Statistics

	Cluster 1 (n = 21)		Cluster 2 (n = 26)	
Variable	**Mean (Sd)**	**Median**	**Mean (Sd)**	**Median**
Score	9.14 (1.20)	10	8.62 (2.10)	10
Time	145 (65)	146	124 (42)	120

4 CONCLUSION

In this study, we identified off-task students by analyzing their reading patterns, however, in terms of academic performance there was no significant difference between off-task and on-task students is noted. In literature, there are many studies found no relationship between off-task behavior and learning outcomes and the reasons for this are not yet known (Cocea et al., 2009; DeFalco, Baker, & D'Mello, 2014). In our case, there might be two possible explanations. First, the quiz sessions students took part in was open book. Therefore, even off-task students might find the answers during the quiz since their average time on quiz higher than students on-task. On the other hand, high-knowledge or high-ability students might also exhibit off-task behaviors since they find the task too easy for them. Simonsen, Little, and Fairbanks (2010) found that less challenging tasks may be less engaging the high-ability students. However, further research is required to test these hypotheses.

Teachers cannot observe and interact with every student at the same time, however, an off-task behavior detector built into the learning environment can observe every student at every moment (Baker, 2007). The obtained results can be used to develop real-time detector for off-task students. Interventions can also be designed to help off-task students (Walonoski & Heffernan, 2006).

ACKNOWLEDGMENT

This research was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (S) Grant Number 16H06304.

REFERENCES

Baker, R. S. (2007). *Modeling and understanding students' off-task behavior in intelligent tutoring systems*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). *Off-task behavior in the cognitive tutor classroom: when students "game the system"*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria.

Cetintas, S., Si, L., Xin, Y. P. P., & Hord, C. (2010). Automatic Detection of Off-Task Behaviors in Intelligent Tutoring Systems with Machine Learning Techniques. *IEEE Transactions on Learning Technologies, 3*(3), 228-236. doi:10.1109/TLT.2009.44

Cocea, M., Hershkovitz, A., & Baker, R. S. J. d. (2009). *The Impact of Off-task and Gaming Behaviors on Learning: Immediate or Aggregate?* Paper presented at the Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling.

DeFalco, J., Baker, R. S. J. d., & D'Mello, S. K. (2014). *Addressing Behavioral Disengagement in Online Learning*.

Flanagan, B., & Ogata, H. (2017). *Integration of Learning Analytics Research and Production Systems While Protecting Privacy.* Paper presented at the 25th International Conference on Computers in Education, ICCE 2017, New Zealand.

Hastie, T., Tibshirani, R., & Walther, G. (2001). Estimating the number of data clusters via the Gap statistic. *J Roy Stat Soc B, 63*, 411-423.

Hofer, M. (2007). Goal conflicts and self-regulation: A new look at pupils' off-task behaviour in the classroom. *Educational Research Review, 2*(1), 28-38. doi:10.1016/j.edurev.2007.02.002

Hughes, K. M. (2010). *Educational Software for Off-Task Behavior*. Unpublished Major Qualifying Project. Worcester Polytechnic Institute, MA.

McElroy-Yeider, & Courtney, Q. (2016). *On-task Behavior During a Reading Task: Effect of Quality Physical Education.* BSU Master's Theses and Projects, Retrieved from http://vc.bridgew.edu/theses/31

Ogata, H., Oi, M., Mohri, K., Okubo, F., Shimada, A., Yamada, M., . . . Hirokawa, S. (2018). Learning Analytics for E-Book-Based Educational Big Data in Higher Education. In H. Yasuura, C.-M. Kyung, Y. Liu, & Y.-L. Lin (Eds.), *Smart Sensors at the IoT Frontier* (pp. 327-350). Cham: Springer International Publishing.

Simonsen, B., Little, C. A., & Fairbanks, S. (2010). Effects of Task Difficulty and Teacher Attention on the Off-Task Behavior of High-Ability Students with Behavior Issues. *Journal for the Education of the Gifted, 34*(2), 245-260.

Walonoski, J. A., & Heffernan, N. T. (2006). *Detection and Analysis of Off-Task Gaming Behavior in Intelligent Tutoring Systems.* Paper presented at the Intelligent Tutoring Systems, Berlin, Heidelberg.